10/9

# 2

OIPE

RAW SEQUENCE LISTING

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

PATENT APPLICATION: US/09/848,852

DATE: 12/18/2001 TIME: 16:52:06

Input Set : N:\Crf3\RULE60\09848852.raw
Output Set: N:\CRF3\12182001\1848852.raw

## ENTERED

## (i) APPLICANT: Hillman, Jennifer L. 5 6 Tang, Y. Tom 7 Corley, Neil C. Guegler, Karl J. 8 Yue, Henry 9 10 Patterson, Chandra (ii) TITLE OF INVENTION: PROTEINS ASSOCIATED WITH EGF SIGNALING 12 (iii) NUMBER OF SEQUENCES: 5 14 (iv) CORRESPONDENCE ADDRESS: 16 17 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc. 18 (B) STREET: 3174 Porter Dr. (C) CITY: Palo Alto 19 20 (D) STATE: CA (E) COUNTRY: USA 21 (F) ZIP: 94304 22 24 (V) COMPUTER READABLE FORM: 25 (A) MEDIUM TYPE: Diskette 26 (B) COMPUTER: IBM Compatible 27 (C) OPERATING SYSTEM: DOS (D) SOFTWARE: FastSEQ for Windows Version 2.0 28 30 (vi) CURRENT APPLICATION DATA: C--> 31 (A) APPLICATION NUMBER: US/09/848,852 C--> 32 (B) FILING DATE: 04-May-2001 34 (vii) PRIOR APPLICATION DATA: 35 (A) APPLICATION NUMBER: 09/069,725 36 (B) FILING DATE: 40 (viii) ATTORNEY/AGENT INFORMATION: 41 (A) NAME: Billings, Lucy J. 42 (B) REGISTRATION NUMBER: 36,749 43 (C) REFERENCE/DOCKET NUMBER: PF-0515 US 45 (ix) TELECOMMUNICATION INFORMATION: 46 (A) TELEPHONE: 650-855-0555 47 (B) TELEFAX: 650-845-4166 (2) INFORMATION FOR SEQ ID NO: 1: 50 52 (i) SEQUENCE CHARACTERISTICS: 53 (A) LENGTH: 298 amino acids 54 (B) TYPE: amino acid 55 (C) STRANDEDNESS: single (D) TOPOLOGY: linear 58 (vii) IMMEDIATE SOURCE: 59 (A) LIBRARY: COLNNOT16 60 (B) CLONE: 1281694 62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1: Met Asn Asp Ser Leu Arg Thr Asp Val Phe Val Arg Phe Gln Pro Glu 65 1 5 10

```
Ser Ile Ala Cys Ala Cys Ile Tyr Leu Ala Ala Arg Thr Leu Glu Ile
67
68
    Pro Leu Pro Asn Arg Pro His Trp Phe Leu Leu Phe Gly Ala Thr Glu
69
                                 40
70
    Glu Glu Ile Gln Glu Ile Cys Leu Lys Ile Leu Gln Leu Tyr Ala Arg
71
                             55
    Lys Lys Val Asp Leu Thr His Leu Glu Gly Glu Val Glu Lys Arg Lys
72
73.
                         70
                                              75
    His Ala Ile Glu Glu Ala Lys Ala Gln Ala Arg Gly Leu Leu Pro Gly
74
75
                     85
                                         90
    Gly Thr Gln Val Leu Asp Gly Thr Ser Gly Phe Ser Pro Ala Pro Lys
76
77
                                     105
78
    Leu Val Glu Ser Pro Lys Glu Gly Lys Gly Ser Lys Pro Ser Pro Leu
79
                                 120
80
    Ser Val Lys Asn Thr Lys Arg Arg Leu Glu Gly Ala Lys Lys Ala Lys
81
                             135
                                                  140
82
    Ala Asp Ser Pro Val Asn Gly Leu Pro Lys Gly Arg Glu Ser Arg Ser
83
                         150
                                             155
    Arg Ser Arg Ser Arg Glu Gln Ser Tyr Ser Arg Ser Pro Ser Arg Ser
85
                                         170
                    165
86
    Ala Ser Pro Lys Arg Arg Lys Ser Asp Ser Gly Ser Thr Ser Gly Gly
87
                                     185
                                                          190
88
    Ser Lys Ser Gln Ser Arg Ser Arg Ser Arg Ser Asp Ser Pro Pro Arg
89
            195
                                 200
                                                      205
90
    Gln Ala Pro Arg Ser Ala Pro Tyr Lys Gly Ser Glu Ile Arg Gly Ser
                             215
    Arg Lys Ser Lys Asp Cys Lys Tyr Pro Gln Lys Pro His Lys Ser Arg
92
93
                         230
                                             235
94
    Ser Arg Ser Ser Ser Arg Ser Arg Ser Arg Glu Arg Ala Asp
95
                    245
                                         250
96
    Asn Pro Gly Lys Tyr Lys Lys Ser His Tyr Tyr Arg Asp Gln Arg
97
                260
                                     265
98
    Arg Glu Arg Ser Arg Ser Tyr Glu Arg Thr Gly Arg Arg Tyr Glu Arg
                                 280
100
     Asp His Pro Gly His Ser Arg His Arg Arg
101
         290
                              295
103 (2) INFORMATION FOR SEQ ID NO: 2:
105
         (i) SEQUENCE CHARACTERISTICS:
106
              (A) LENGTH: 3464 base pairs
107
              (B) TYPE: nucleic acid
108
              (C) STRANDEDNESS: single
              (D) TOPOLOGY: linear
109
       (vii) IMMEDIATE SOURCE:
111
112
              (A) LIBRARY: COLNNOT16
              (B) CLONE: 1281694
113
115
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
     TTCGTGAAGC ACTCCATGGA GCATGTGTCA ATGGCCTGTG TCCACCTGGC TTCCAAGATA
                                                                             60
117
118
     GAAGAGGCCC CAAGACGCAT ACGGGACGTC ATCAATGTGT TTCACCGCCT TCGACAGCTG
                                                                            120
     AGAGACAAAA AATAATCGTT ATGTACCTTC AGGTGTTAGA GTGTGAGCGT AACCAACACC
119
                                                                            180
```

120	TGGTCCAGAC	CTCATGGGTA	GCCTCTGAGG	GTAAGTGACT	AAGACTTCTC	CTCTGCTGTC	240
121	CAAGCGCTTT	GGTGCAGGGA	CAGCGGCATC	TTCAGCCAAT	CCAGTGCAGG	CTCTCCACCG	300
122	AAGGCTGGCT	CTAGACTGGT	GACCCCTTGT	TGAAATGGGA	CAGTTGGCAG	CGGCTCTGAT	360
123	GAGCCCGAGA	AGAGGCCTGC	CCTTGGGTGC	GGAGTCTCCC	TCCGCACGAT	GCTCCCACGC	420
124	GTCCAACTTG	CACCCAAGGG	GCTTTTCCCT	CTTCCAAGTG	GACTCCTTCA	AGGAAGCTGC	480
125	AGCTCGGTCA	GCAGAGAAGG	GGCCTGCCGC	CAGCGCCCTG	GAGGAAGAGG	AAGAGGAACC	540
126	CAAGAGGATG	GCTTGTCTCC	CAGCAGCCAC	ACCGGCTTTG	TGCTCAGCCA	GTTCATTTGA	600
127	GTTTGCATGT	TTCTCTGCAC	TATGGATTTT	GAGCATTTAG	ATTTCTTTAA	TCAAAAGCGT	660
128	TTTAGTGACT	CCAGTAGACA	TTTTCTTTCT	GAGGCATCGT	GCTTTGCATG	AGAGCAGGCC	720
129	AAGGTTGAGG	GGAAAAGTAA	AGTTAAAGTC	GGTTCTCTTT	CATAGCAACA	CGTATTGTCT	780
130	GACATTCAGC	CAGCTTTTTT	TTTTTCTAAT	AATTTCTGTG	CCTTTCTGTC	CTGTATTTAC	840
131	TGTATTTAGA	AAAAGCAGCT	AGAATATTTC	TCCATTAACT	CTTGAGATTC	ACAGGACTGT	900
132	CTAGCTCTGA	GTCCTAGCAA	TAGACTCCTT	AGAGGAGTAG	TACGTTTATC	TAGATTTTCT	960
133	CTAGATAATG	CAGGCGGAAG	ACCTGGGTTC	CCGGGTGGGG	CATTGCAGTT	CTTCCTGTGT	1020
134	TTGGCTTCCA	GGAATTACAT	GAACGACAGC	CTTCGCACCG	ACGTCTTCGT	GCGGTTCCAG	1080
135	CCAGAGAGCA	TCGCCTGTGC	CTGCATTTAT	CTTGCTGCCC	GGACGCTGGA	GATCCCTTTG	1140
136	CCCAATCGTC	CCCATTGGTT	TCTTTTGTTT	GGAGCAACTG	AAGAAGAAAT	TCAGGAAATC	1200
137	TGCTTAAAGA	TCTTGCAGCT	TTATGCTCGG	AAAAAGGTTG	ATCTCACACA	CCTGGAGGGT	1260
138			CGCTATCGAA				1320
139	CCTGGGGGCA	CACAGGTGCT	GGATGGTACC	TCGGGGTTCT	CTCCTGCCCC	CAAGCTGGTG	1380
140	GAATCCCCCA	AAGAAGGTAA	AGGGAGCAAG	CCTTCCCCAC	TGTCTGTGAA	GAACACCAAG	1440
141	AGGAGGCTGG	AGGGCGCCAA	GAAAGCCAAG	GCGGACAGCC	CCGTGAACGG	CTTGCCAAAG	1500
142	GGGCGAGAGA	GTCGGAGTCG	GAGCCGGAGC	CGTGAGCAGA	GCTACTCGAG	GTCCCCATCC	1560
143	CGATCAGCGT	CTCCTAAGAG	GAGGAAAAGT	GACAGCGGCT	CCACATCTGG	TGGGTCCAAG	1620
144			CAGGAGTGAC				1680
145	CCCTACAAAG	GCTCTGAGAT	TCGGGGCTCC	CGGAAGTCCA	AGGACTGCAA	GTACCCCCAG	1740
146	AAGCCACACA	AGTCTCGGAG	CCGGAGTTCT	TCCCGTTCTC	GAAGCAGGTC	ACGGGAGCGG	1800
147	GCGGATAATC	CGGGAAAATA	CAAGAAGAAA	AGTCATTACT	ACAGAGATCA	GCGACGAGAG	1860
148	CGCTCGAGGT	CGTATGAACG	CACAGGCCGT	CGCTATGAGC	GGGACCACCC	TGGGCACAGC	1920
149	AGGCATCGGA	GGTGAGGCGG	GGTTGCAGTG	ACTGGTGGCC	GCAAGCCCTT	CCCTGGGGAG	1980
150	TACCTGATGG	CTGCCCTTTG	ACCCCCGGTG	GCTGCCCTTT	GACCCCCGGG	TGTGCTCTCA	2040
151	GCGCAAGTGG	TCCTAGAACA	GGATTCTTTT	TGGAAATGTC	TGTCGACTGG	ACCTTGGTGG	2100
152			GACCGGTGAC				2160
153			GCTCAGCTCG				2220
154			TTGGCGTGGC				2280
155	TCTCCACTCC	CCGACTGATA	CTCAATTACG	TGAAGCCAAG	AAAGATGATT	TTTAGAACCT	2340
156			TTATGTACAT				2400
157	CTTAACTGCC	CCTTATTCTC	TCTCCACGTT	GTAAATAAAC	ATGTGTTTAA	TACAAGTTAA	2460
158			AACTTGAATC				2520
159	TGACTTTTAA	AATGTGAGGG	TATTTGGATC	TGTGTTGAAA	GTCGTATATT	TTTATCTGTG	2580
160			CAGCTCCTAA				2640
161			TCCACTCAGG				2700
162			AAGTGTCCAG				2760
163			ACTGAAGTTT				2820
164			GGGGCCTGAG				2880
165			AAGTTTTGGA				2940
166			AACCCATGAC				3000
167			CACGGGGGTG				3060
168			GACACAGTTT				3120

```
169 GGTGGTACCG TTTGGACTTA CTAGGGCAGT GGGACATATA GGCCGGGGCT AGTGGGATAA
                                                                          3180
170 CGGGGAGTTA CGCCTGATGA CTTTTTTGAT GGAATCCTGC ATTAGATAGC TGGTGGGACC
                                                                          3240
171 CCCCCCTCAG AATTGGGGAA CTGAGGAGAC TCCAGGGAGG GTGTCCTTCC AGGGAGAGCA
                                                                          3300
172 GCTATGAGGG GCCCCCTAGC TTCCTGTGCC TGGAAGTAAG AGAACCAGTA AAGGGCCATA
                                                                          3360
173 CACACCTGTA CCCAAGAGAC CGCTCTCCAT TTGCTTTCTT TTTTTACTAA ATAATTGTAA
                                                                          3420
     AATATTATTA TGACATAAAG AACCATTTAA GGCCAAAAAA AAAA
                                                                          3464
176 (2) INFORMATION FOR SEQ ID NO: 3:
178
         (i) SEQUENCE CHARACTERISTICS:
179
              (A) LENGTH: 332 amino acids
180
              (B) TYPE: amino acid
181
              (C) STRANDEDNESS: single
182
              (D) TOPOLOGY: linear
184
       (vii) IMMEDIATE SOURCE:
1.85
              (A) LIBRARY: BEPINOT01
186
              (B) CLONE: 2056178
188
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
     Met Tyr Ser Ala Gln Arg Phe Trp Gly Thr Ile Trp Ala Arg Arg Gly
190
191
                      5
                                          10
192
     Ala His Leu Ala Pro Pro Asp Ala Ser Ile Leu Ile Ser Asn Val Cys
193
                                      25
194
     Ser Ile Gly Asp His Val Ala Gln Glu Leu Phe Gln Gly Ser Asp Leu
195
                                  40
196
     Gly Met Ala Glu Glu Ala Glu Arg Pro Gly Glu Lys Ala Gly Gln His
197
                              55
198
     Ser Pro Leu Arg Glu Glu His Val Thr Cys Val Gln Ser Ile Leu Asp
199
                         70
                                              75
200
     Glu Phe Leu Gln Thr Tyr Gly Ser Leu Ile Pro Leu Ser Thr Asp Glu
201
202
     Val Val Glu Lys Leu Glu Asp Ile Phe Gln Gln Glu Phe Ser Thr Pro
203
                 100
                                      105
204
     Ser Arg Lys Gly Leu Val Leu Gln Leu Ile Gln Ser Tyr Gln Arg Met
205
                                 120
                                                      125
206
     Pro Gly Asn Ala Met Val Arg Gly Phe Arg Val Ala Tyr Lys Arg His
207
                             135
208
     Val Leu Thr Met Asp Asp Leu Gly Thr Leu Tyr Gly Gln Asn Trp Leu
209
                         150
                                              155
210
     Asn Asp Gln Val Met Asn Met Tyr Gly Asp Leu Val Met Asp Thr Val
211
                     165
                                          170
212
     Pro Glu Lys Val His Phe Phe Asn Ser Phe Phe Tyr Asp Lys Leu Arg
213
                 180
                                     185
214
     Thr Lys Gly Tyr Asp Gly Val Lys Arg Trp Thr Lys Asn Val Asp Ile
215
                                 200
216
     Phe Asn Lys Glu Leu Leu Ile Pro Ile His Leu Glu Val His Trp
217
                             215
218
     Ser Leu Ile Ser Val Asp Val Arg Arg Thr Ile Thr Tyr Phe Asp
219
                         230
                                              235
     Ser Gln Arg Thr Leu Asn Arg Arg Cys Pro Lys His Ile Ala Lys Tyr
220
221
                     245
                                          250
222
     Leu Gln Ala Glu Ala Val Lys Lys Asp Arg Leu Asp Phe His Gln Gly
```

```
223
                 260
                                      265
                                                          270
224
     Trp Lys Gly Tyr Phe Lys Met Asn Val Ala Arg Gln Asn Asn Asp Ser
225
             275
                                  280
                                                      285
226
     Asp Cys Gly Ala Phe Val Leu Gln Tyr Cys Lys His Leu Ala Leu Ser
227
                             295
                                                  300
     Gln Pro Phe Ser Phe Thr Gln Gln Asp Met Pro Lys Leu Arg Arg Gln
228
229
                         310
                                              315
230
     Ile Tyr Lys Glu Leu Cys His Cys Lys Leu Thr Val
231
                     325
                                          330
233 (2) INFORMATION FOR SEQ ID NO: 4:
235
         (i) SEQUENCE CHARACTERISTICS:
236
              (A) LENGTH: 1991 base pairs
237
              (B) TYPE: nucleic acid
238
              (C) STRANDEDNESS: single
239
              (D) TOPOLOGY: linear
241
       (vii) IMMEDIATE SOURCE:
242
              (A) LIBRARY: BEPINOT01
243
              (B) CLONE: 2056178
245
        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
247
     GCCTCCCTGT CCCCCGACCC TCTTTTGATG CCTCAGCAAG TGAAGAGGAG GAAGAAGAGG
                                                                            60
     AGGAGGAGGA GGATGAAGAT GAAGAGGAGG AAGTGGCAGC TTGGAGGCTG CCCCCAAGAT
                                                                           120
249
     GGAGTCAGCT GGGAACCTCC CAGCGGCCCC GCCCTTCCCG CCCCACTCAT CGAAAAACCT
                                                                           180
250
    GCTCACAGCG CCGCCGCA GCCATGAGAG CCTTCCGGAT GCTGCTCTAC TCAAAAAGCA
                                                                           240
     CCTCGCTGAC ATTCCACTGG AAGCTTTGGG GGCGCCACCG GGGCCGGCGG CGGGGCCTCG
                                                                           300
252
    CACACCCCAA GAACCATCTT TCACCCCAGC AAGGGGGTGC GACGCCACAG GTGCCATCCC
                                                                           360
    CCTGTTGTCG TTTTGACTCC CCCCGGGGGC CACCTCCACC CCGGCTGGGT CTGCTAGGTG
                                                                           420
    CTCTCATGGC TGAGGATGGG GTGAGAGGGT CTCCACCAGT GCCCTCTGGG CCCCCCATGG
                                                                           480
    AGGAAGATGG ACTCAGGTGG ACTCCAAAGT CTCCTCTGGA CCCTGACTCG GGCCTCCTTT
                                                                           540
256
    CATGTACTCT GCCCAACGGT TTTGGGGGGAC AATCTGGGCC AGAAGGGGAG CGCACTTGGC
                                                                           600
257
     ACCCCTGAT GCCAGCATCC TCATCAGCAA TGTGTGCAGC ATCGGGGACC ATGTGGCCCA
                                                                           660
258
     GGAGCTTTTT CAGGGCTCAG ATTTGGGCAT GGCAGAAGAG GCAGAGAGGC CTGGGGAGAA
                                                                           720
259
     AGCCGGCCAG CACAGCCCCC TGCGAGAGGA GCATGTGACC TGCGTACAGA GCATCTTGGA
                                                                           780
    CGAATTCCTT CAAACGTATG GCAGCCTCAT ACCCCTCAGC ACTGATGAGG TAGTAGAGAA
260
                                                                           840
261
     GCTGGAGGAC ATTTTCCAGC AGGAGTTTTC CACCCCTTCC AGGAAGGGCC TGGTGTTGCA
                                                                           900
     GCTGATCCAG TCTTACCAGC GGATGCCAGG CAATGCCATG GTGAGGGGCCT TCCGAGTGGC
                                                                           960
    TTATAAGCGG CACGTGCTGA CCATGGATGA CTTGGGGGACC TTGTATGGAC AGAACTGGCT
263
                                                                          1020
    CAATGACCAG GTGATGAACA TGTATGGAGA CCTGGTCATG GACACAGTCC CTGAAAAGGT
264
                                                                          1080
265
    GCATTTCTTC AATAGTTTCT TCTATGATAA ACTCCGTACC AAGGGTTATG ATGGGGTGAA
                                                                          1140
266
    AAGGTGGACC AAAAACGTGG ACATCTTCAA TAAGGAGCTA CTGCTAATCC CCATCCACCT
                                                                          1200
267
    GGAGGTGCAT TGGTCCCTCA TCTCTGTTGA TGTGAGGCGA CGCACCATCA CCTATTTTGA
                                                                          1260
268
    CTCGCAGCGT ACCCTAAACC GCCGCTGCCC TAAGCATATT GCCAAGTATC TACAGGCAGA
269
    GGCGGTAAAG AAAGACCGAC TGGATTTCCA CCAGGGCTGG AAAGGTTACT TCAAAATGAA
270
    TGTGGCCAGG CAGAATAATG ACAGTGACTG TGGTGCTTTT GTGTTGCAGT ACTGCAAGCA
                                                                          1440
271
    TCTGGCCCTG TCTCAGCCAT TCAGCTTCAC CCAGCAGGAC ATGCCCAAAC TTCGTCGGCA
                                                                          1500
272
    GATCTACAAG GAGCTGTGTC ACTGCAAACT CACTGTGTGA GCCTCGTACC CCAGACCCCA
                                                                          1560
273
    AGCCCATAAA TGGGAAGGGA GACATGGGAG TCCCTTCCCA AGAAACTCCA GTTCCTTTCC
                                                                          1620
274
    TCTCTTGCCT CTTCCCACTC ACTTCCCTTT GGTTTTTCAT ATTTAAATGT TTCAATTTCT
                                                                          1680
275
    GTATTTTTT TTCTTTGAGA GAATACTTGT TGATTTCTGA TGTGCAGGGG GTGGCTACAG
                                                                          1740
276
    AAAAGCCCCT TTCTTCCTCT GTTTGCAGGG GAGTGTGGCC CTGTGGCCTG GGTGGAGCAG
                                                                          1800
```

VERIFICATION SUMMARY

DATE: 12/18/2001 TIME: 16:52:07

PATENT APPLICATION: US/09/848,852

Input Set : N:\Crf3\RULE60\09848852.raw
Output Set: N:\CRF3\12182001\1848852.raw

L:31 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:] L:32 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]